

AMCoR

Asahikawa Medical College Repository <http://amcor.asahikawa-med.ac.jp/>

北海道放射線技術雑誌 (1984.08) 43～44号:101～104.

4係数Logistic関数の収束性の拡張並びに応用回帰分析

西部茂美、石川幸雄

4 係数 Logistic 関数の収束性の拡張並びに 応用回帰分析

旭川医科大学医学部附属病院放射線部

◎西部 茂美 石川 幸雄

はじめに

ある種の¹⁾非線型の問題においては、正規方程式を書き下して反復法を用いて解くのが最も便利である。この方法がうまくゆくかどうかは、方程式の型と用いる反復法とによって決まる。この中には、代表的なものとして、a. 線型化法 b. 最急勾配法 c. Marguardt の中間法、等があるが、ここでは a の方式で解くことにする。

(1) 線型化法 (Linearization method) あるいは テーラー展開法

各段階で線型最小二乗法によって得られた結果を利用する方法で、今仮定する模型を²⁾ $y = (a, b, c, d)$

$$= \frac{a-b}{1 + (\frac{x}{c})^b} + d \dots\dots\dots(1) \text{式と定め、求めるべき}$$

a, b, c, d の真の値は、近似値を a_0, b_0, c_0, d_0 又、誤差を $\Delta a, \Delta b, \Delta c, \Delta d$ とすると $a = a_0 + \Delta a, b = b_0 + \Delta b, c = c_0 + \Delta c, d = d_0 + \Delta d$ となり、要するに $\Delta a, \Delta b, \Delta c, \Delta d$ を求めれば a, b, c, d は求まる。この関数の第1次導関数までの³⁾テーラー展開は、 a_0, b_0, c_0, d_0 について、

$$y = f(a_0, b_0, c_0, d_0) + \Delta a \left(\frac{\partial f}{\partial a} \right) + \Delta b \left(\frac{\partial f}{\partial b} \right) + \Delta c \left(\frac{\partial f}{\partial c} \right) + \Delta d \left(\frac{\partial f}{\partial d} \right) \dots\dots\dots(2) \text{式となり、}$$

各母数 a, b, c, d について³⁾偏微分すると、 $\frac{\partial f}{\partial a} =$

$$\frac{\partial f}{\partial a} = \frac{1}{1 + (\frac{x}{c})^b} = P \dots\dots\dots(3)$$

$$\frac{\partial f}{\partial b} = (d-a) \cdot \ln \left(\frac{x}{c} \right) \cdot \left(\frac{x}{c} \right)^b \cdot P^2 \dots\dots\dots(4)$$

$$\frac{\partial f}{\partial c} = \frac{b \cdot (a-d)}{c} \cdot \left(\frac{x}{c} \right)^b \cdot P^2 \dots\dots\dots(5)$$

$$\frac{\partial f}{\partial d} = 1 - P \dots\dots\dots(6)$$

(3)~(6)式で表わされ、従がって、これらの式を整理すると、 $y = f(a_0, b_0, c_0, d_0) + \Delta a \cdot P + \Delta b \cdot (d-a) \cdot \ln \left(\frac{x}{c} \right) \cdot \left(\frac{x}{c} \right)^b \cdot P^2 + \Delta c \cdot \frac{b \cdot (a-d)}{c} \cdot \left(\frac{x}{c} \right)^b \cdot P^2 + \Delta d \cdot (1-P) \dots\dots\dots(7) \text{式となり、}$

この式は $\Delta a, \Delta b, \Delta c, \Delta d$ に関して線型であり、直線回帰と同じ考え方を適用出来る。我々は、非線型模型と与えられたデータに対し、誤差二乗和を次のよ

$$\begin{bmatrix} \sum (\frac{\partial f}{\partial a})^2, \sum \frac{\partial f}{\partial a} \frac{\partial f}{\partial b}, \sum \frac{\partial f}{\partial a} \frac{\partial f}{\partial c}, \sum \frac{\partial f}{\partial a} \frac{\partial f}{\partial d} \\ \sum \frac{\partial f}{\partial b} \frac{\partial f}{\partial a}, \sum (\frac{\partial f}{\partial b})^2, \sum \frac{\partial f}{\partial b} \frac{\partial f}{\partial c}, \sum \frac{\partial f}{\partial b} \frac{\partial f}{\partial d} \\ \sum \frac{\partial f}{\partial c} \frac{\partial f}{\partial a}, \sum \frac{\partial f}{\partial c} \frac{\partial f}{\partial b}, \sum (\frac{\partial f}{\partial c})^2, \sum \frac{\partial f}{\partial c} \frac{\partial f}{\partial d} \\ \sum \frac{\partial f}{\partial d} \frac{\partial f}{\partial a}, \sum \frac{\partial f}{\partial d} \frac{\partial f}{\partial b}, \sum \frac{\partial f}{\partial d} \frac{\partial f}{\partial c}, \sum (\frac{\partial f}{\partial d})^2 \end{bmatrix} \begin{bmatrix} \Delta a \\ \Delta b \\ \Delta c \\ \Delta d \end{bmatrix} = \begin{bmatrix} \sum \frac{\partial f}{\partial a} (Y_k - f(a_0, b_0, c_0, d_0)) \\ \sum \frac{\partial f}{\partial b} (Y_k - f(a_0, b_0, c_0, d_0)) \\ \sum \frac{\partial f}{\partial c} (Y_k - f(a_0, b_0, c_0, d_0)) \\ \sum \frac{\partial f}{\partial d} (Y_k - f(a_0, b_0, c_0, d_0)) \end{bmatrix} \dots\dots\dots(9)$$

Fig. 1 ---(9)

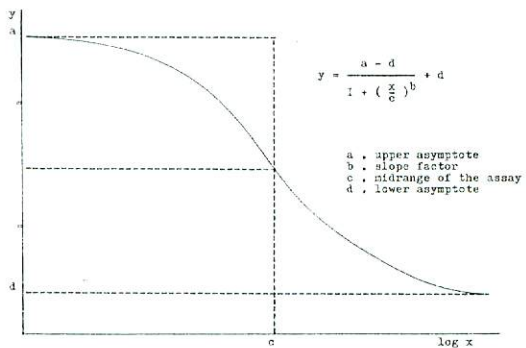


Fig. 2

うに定義する。

$$S(a, b, c, d) = \sum_{k=1}^n \left\{ Y_k - f_k(a_0, b_0, c_0, d_0) \right\}$$

$$-\Delta a \left(\frac{2f_k}{2a} \right) - \Delta b \left(\frac{2f_k}{2b} \right) - \Delta c \left(\frac{2f_k}{2c} \right) - \Delta d \left(\frac{2f_k}{2d} \right) \dots\dots\dots(8)$$

(8)式を $\Delta a, \Delta b, \Delta c, \Delta d$ で偏微分し、最小値を求めるための正規方程式を作る。この連立方程式を解き、求めた $\Delta a, \Delta b, \Delta c, \Delta d$ より新たな a, b, c, d を計算し、最初に戻り、 $\Delta a, \Delta b, \Delta c, \Delta d$ は次第に0収束してゆくの、ある程度小さくなったら計算を打ち切る。この正規方程式を解くために、Fig. 1で表わす4行4列の行列式を考える。⁴⁾Gauss-Jordanの消去法により、 $\Delta a, \Delta b, \Delta c, \Delta d$ を得る。尚初期値となるグラフ上の意味は一般的にFig. 2のごとくである。すなわち、 y をresponse、 x をdoseとすると、係数 a は B_0 に相当する上の漸近線、 b はslope factorで理想的には1とし、 c は50% interceptに相当する。

反復手順は、解が収束つまり相続く反復により、すなわち $a_k = a_{0k} + \Delta a_k, b_k = b_{0k} + \Delta b_k$ 、以下 c, d についても同様に k と $(k+1)$ 番目において、

$$\left| \frac{a_{k+1} - a_k}{a_k} \right| < \epsilon \text{ が成立するまで続けられる。ここ}$$

にはあらかじめ決められた量で0.0001とした。しかしこのままでは、RIAのスタンダード・カーブによっては、収束を期待出来ないものがかかなり多くみられ、これらの初期値は、得られた情報に基づく推測又は予備的な推定によるものであってよいので、係数 a をグラフ上の値の1.2倍、 c は2番目の観測値の y responseと x doseを用いて、 $C = \frac{x_2 \cdot (y_2 - d)}{(a - y_2)}$ とし、 x doseの両端点を除きどの値を用いてもあまり問題にはならず、 d をグラフ上の値の0.8倍とすることにより、収束性は良好であった。

尚、線型化の手続きは、ある種のデータに対して以下の欠点を持っている。

- a. 収束度はきわめて遅く、1回の収束時間は、YHP-Model 30処理で、30秒はかかる。
 - b. たえず広範囲にわたって振動することがあり、時として平方和の減少と同様増大も起こることがあるが、解は安定する。
 - c. 全く収束しないどころか発散さえするために、平方和は反復ごとに限りなく増大することがある。
- それでこれらの欠点に対して、誤差二乗和を $S(a_{k+1})$

$> S(a_k)$ なら Δa_k を $\frac{1}{2}$ とし、もし $S(a_{k-1}) < S(a_k)$ なら2倍とするように修正を行なった。この $\frac{1}{2}$ 及び2倍の手続きは、 a_k 及び a_{k+1} の間3点だが、その間に $S(a)$ の局所的最小値を含むようになるまで続け、最小位置の決定には2次内挿を用い、再び反復を開始する。

(2) 回帰模型の詳細な解析

線型模型に適する回帰線の推定に対する精密度の測定として、観測値の予測された、又は当てはめられた値との残差平方和と、繰り返し測定による純誤差平方和を求め、両者の差により模型不適合平方和を求めて、それらの自由度で除することにより、純誤差不偏分散ならびに、模型不適合による不偏分散を求め、その比よりF分布95%点と比較することにより、推定模型が有意であるかないかを一つの判定として用いられている。しかし非線型模型に対しては一般的にいって適

```

E0 <CPI>
      9316
      8999
      C.V. = 2.289321907 %
NO. ST.D-H0D0      B-E0 (%)      COUNT*1      COUNT*2      C.V.(%)
1 1 93.24193265      0542      6442      0.592759
2 2 82.85296651      7914      7790      0.24053
3 3 65.0066248      5975      5866      0.42053
4 4 41.32308537      3752      3772      0.49556
5 5 24.17279798      2223      2162      0.48336
6 6 15.09744716      1406      1344      0.5254

*****CALCULATION*****

*RE-CL*
*RE-CL*
*RE-CL*
*RE-CL*

Y=A-D/(1+(X-C)/B)+E

INITIAL VALUE      A      B      C      D
FINAL VALUE      99.28512391      0.905169581      14.6134607      4.86922570
FINAL VARIANCE      4.12532E-05      -1.30364E-06      -1.5923E-05      -1.96005E-06

NO. ST.D-H0D0      *B-E0 (%)
1 1 93.31335536
2 2 83.73165578
3 3 65.03507336
4 4 41.44474539
5 5 23.97223267
6 6 15.26400738

R = 0.999996181      *B(SLOPE.FACTOR) = 0.985169581

*****ANALYSIS OF VARIANCE TABLE*****
SV      SS      DF      MS      F
CENTRAL      F(0.05) = 5.140
TABLE      F(0.05) = 24.1731
ZHEH      5.506      8.000      ST2 = 0.688      0.090 (NS)
FUTEKIGOU      0.179      2.000      HSL = 0.089      0.131 (F.T)
JUNGOSH      5.327      6.000      SE2 = 0.888      0.120 (NF)

* HSL = 0.321      IF HSL > 2.2      THEN RUN HSSAYS AGAIN **

```

Fig. 3

当であるとはいえない。実際的な方法として、まだ説明されない変動 $V(Y_k) = b^2$ の推定値と比較することはできるが、任意の水準での結論を得るために、^{5)F}

統計量を使用することはできない。

Fig. 3はTSHのデータ処理の1例であり、分散分析表中のSV, SS, DF, MS, Fは、それぞれ変動要因 (source of variation), 平方和 (sum of square), 自由度 (degrees of freedom), 不偏分散 (mean of square), 5%危険率におけるF値である。又、回帰曲線の適合度の判定として、⁶⁾ 大阪大学市原らが提唱している方法で、指標NSは回帰曲線のまわりの総合的な誤差の標準化値を、NSLは回帰不適合誤差の標準化値を、NFは回帰不適合誤差の純誤差に対する比率の標準化値を示し、従来から当院でも回帰適合度の1つの判定としてきた。しかし、ここでまだ説明されない変動誤差二乗和 $S(a, b, c, d) = 5.506$ である。非線型の場合に対する厳密な結果を欠くときは、この二乗和を近似的に自由度ここでは $12 - 4 = 8$ (4つの母数は既に推定されているので) における値とみなす。非線型の場合、模型が正しくてもこれを線型の場合のように b^2 の不偏推定量とみなすわけにはいかない。 b^2 の純誤差の推定量は、繰り返し測定によって得られる。これは今の場合、各スタンダード・ドーズを duplicate で測定しているの、自由度6の二乗和は 0.197 (自由度 $8 - 6 = 2$) によって与えられ、次の不

$$\text{不偏分散} \frac{0.197}{2} = 0.089, \frac{5.327}{6} = 0.888 \text{ を比較する。}$$

今の場合F検定は適用できないが、比較の測度として、

$$F(2, 6, 0.95) = 5.140 \text{ を使用する。} \frac{0.089}{0.888} = 0.101$$

という値を見ると、模型がそれほど不適合であるとも考えられない。市原は、NSLが2.2を越えると模型不適合としているが、NSLが2.2以下もF検定値も大きく上まわる場合があり、我々はF検定値の $\frac{1}{3}$ の値との比較により radio assays の適合度の目安としている。このことは、さらに観測値と回帰データとの相関が0.9995以上を示すものである。

(3) Spline関数との比較

これは、検量線をいくつか分割し、各分割区間に多項式をあてはめ、分割点を連続に保つので、検量線の特性によらず良好な回帰が得られる点で、最近応用されつつある。

Fig. 4は、さきほどのTSHのデータに関し、3次多項式スプライン関数で処理をしたものである。図表の分散分析表中からも解るように、繰り返し測定によ

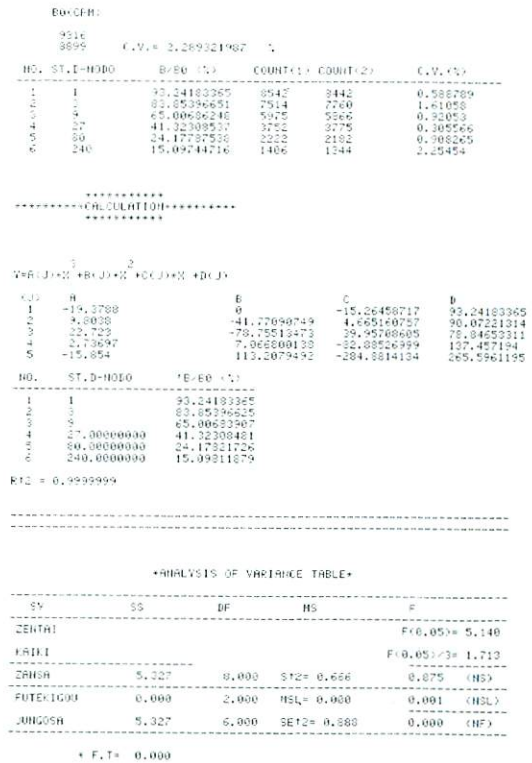


Fig. 4

る純誤差平方和が残差平方和と同一であり、模型不適合平方和は0に等しい。という事は spline 関数で処理すれば模型不適合による不偏分散が限りなく0に近づくけれども、飛び離れた観測点があったり、systematicな検量線の偏位がある場合には問題となる。そこで我々は、4係数 logistic 関数で処理出来ないような、例えば、ある種の Free T₄ のスタンダード・カーブや、4係数 logistic 関数処理で現データとの相関が0.9999以上のものに関して Spline 処理を行なっている。

(4) 結 語

純誤差不偏分散の値が0.8以下小さく、しかも相関が0.9997と良い場合でも純誤差平方和が模型不適合平方和より小さい場合は、観測データと推定模型との関係を見るに、assays がきわめて適切に行なわれたとは判断しがたい場合がある。というのは、観測データの繰り返しの誤差が小さくすなわち、ピベッティングの誤差が小さいような場合でも、観測データが本来の確からしい適切濃度を示さない y の response を kit の

標準濃度の内、1個もしくは2個ある場合に考えられ、各 assays の傾向を調べる上できわめて重要で、希釈して標準濃度を作成する場合においても重要になってくる。又純誤差平方和が10の値を越え、しかも相関が0.99と悪く、尚かつ模型不適合平方和が小さい場合で、F統計量の $\frac{1}{3}$ の値を越えないこともある。これは、繰り返しによる誤差がきわめて大きい場合で、模型不適合不偏分散は有意ではない。このデータを推定模型に回帰させるのに無理があるからで、我々は1つの指標として、純誤差平方和を10以下に保つように心がけ、かつF統計量の $\frac{1}{3}$ の値で assays の傾向を調べ、NSLが2.2以下の値を採用している。

参考文献

- 1) Draper NR, Smith H : 応用回帰分析, 中村慶一訳, 森北出版, 東京, 1967
- 2) Rodbard D, Hutt DM ; Statistical analysis of radioimmunoassays and immunoradiometric assays. A generalized, weighted, iterative, least-squares method for logistic curve fitting in Symposium on RIA and Related Procedures in Medicine, Int Atomic Energy Agency, Vienna, 1974, P 165.
- 3) 矢野健太郎 石原 繁 著 科学技術者のための基礎数学, 裳華房
- 4) 戸川隼人 計算機のための数値計算, コンピュータテキスト5 サイエンス社
- 5) 小島治夫 簡約統計学, 南江堂, 1965
- 6) 市原清志 ラジオイムノアッセイの数量的解析とデータ処理, 日本放射線技術学会雑誌35巻3号, 364 (1979)